DOCUMENT RESUME

ED 138 627 TM 006 250

AUTHOR Bart, William M.; Lele, Kaustubh
TITLE Defusing the Intelligence x Race Debate: Comparison of Intelligence Test Item Hierarchies for Two Races
PUB DATE [Apr 77]
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Black Students; *Caucasian Students; Comparative Testing; Elementary Education; Elementary School Students; Intelligence Differences; *Intelligence Tests; *Item Analysis; Racial Differences; *Twins
IDENTIFIERS *Item Hierarchies; Ordering Theory; Ravens Progressive Matrices

ABSTRACT

One hundred eighty one sets of black twins and 223 sets of white twins provided responses to four 12-item subtests of the Raven's Progressive Matrices Test, Standard Version. The children were in elementary school and their item response patterns were analyzed with the use of revised ordering-theoretic methods to search for best-fitting hierarchies without any item intransitivities occurring. The item hierarchies generated for the two races for each subtest were found to be quite similar in terms of the number of item prerequisite relations that were shared. However, some items were located as acting quite differently with respect to their placements in their respective item hierarchies. One benefit to item hierarchy analysis and ordering theory is that the learning hierarchy and/or pattern of acquisition for a set of test items for each of several social groups can be easily determined. (Author/RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED138627

Session D-17

Defusing the Intelligence x Race Debate: Comparison of Intelligence Test Item Hierarchies for Two Races

by

William M. Bart and Kaustubh Lele
University of Minnesota

Paper presented at the Annual Meeting of
the American Educational Research Association,
New York, 1977

TM006 250

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

WILLIAM M. BART

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

How can racial differences in intellectual performance be described in a positive manner so that 1) steps for subsequent instruction to improve intellectual performance can be recommended, 2) patterns of intellectual skill acquisition can be discussed, and comparative statements that lead to charges of "racism," "racial superiority," and "prejudice" can be avoided? This paper is inteded to offer an answer to that question as some of the benefits of item hierarchy analysis for differential and educational psychology are described. Procedures in the analysis of item hierarchies have already demonstrated their utility in validating a priori instructional sequences and learning hierarchies (Airasian and Bart, 1975), in suggesting sequences of instructional activities (Partin, 1976), and in indentifying patterns of intellectual skill acquisition (Bart and Airasian, 1974). In addition, item hierarchies are also especially appropriate in describing group differences without eliciting charges of "group superiority," "prejudice," etc.

To explicate that last propostion consider two cases. First, let us assume that total test mean scores are used to compare groups. Such scores as well as arithmetical derivatives of the scores such as z - scores can be linearly ordered with the basic relation of "greater than or equal to" that is a reflexive, assymmetric, transitive, and connected relation being used to interrelate the scores. The "$\geq$" relation which would interrelate two total test mean scores can readily be translated into the evaluative relation of "better than or equal to" which would interrelate the two corresponding groups. Thus, the statement that total intelligence test mean score of 40 for group A is greater than the total intelligence test mean score of 30 for group B is readily translated into the statement that group A is better or has more intelligent individuals than group B which, in turn, triggers emotive

counter charges. What is then being argued is that much of the controversy about the relationship of intelligence to race is derived from the overusage of group measures that are linearly ordered. The type of measure used to compare groups does have impact on the types of evaluative statements subsequently made about the groups.

Secondly, let us assume that test item hierarchies are used to compare groups. Such measures are not linearly ordered but are partially ordered with the basic relation of "includes" that is reflexive assymmetric, transitive and unconnected. The "includes" relation can be used to partially order item hierarchies but cannot be translated into the evaluative relation of "better than or equal to" which would interrelate the corresponding groups. Thus, to say that group A engenders item hierarchy 1 and that group B engenders item hierarchy 2 for the same intelligence test items tells us nothing about whether group A is smarter than group B or vice versa. Comparison of group item hierarchies can only lead to the identification of similarities and differences among the patterns of item process acquisition for the two groups. To describe group differences in terms of item hierarchies precludes evaluative statements of group superiority and counter charges of "racism," "prejudice," etc., because item hierarchies are partially ordered and most controversy-producing evaluative statements are based on the usage of linear ordering relations.

The concepts, principles and procedures of item hierarchies allow for the integration of instruction, measurement, and developmental psychology, so that, for example, groups can be described and compared in a positive, constructive manner with educational import and without the negative, socially-divisive overtones that have afflicted so many discussions of racial differences in intelligence (Block and Dworkin, 1976). In this paper, the item hierarchies for subtests of the Raven's Progressive Matrices test for black and white children will be described, compared and discussed.

Method

Subjects

The subjects were 181 sets of black twins and 223 sets of white twins. All of these twins were elementary school children and the data in this study were previously collected by Dr. Sandra Scarr and her colleagues.

Test

The 808(=362 + 446) subjects were administered the Standard Version of the Raven's Progressive Matrices Test which is widely accepted as a culture-fair measure of general intelligence, "g." The response patterns of each subject to each of the first four 12-item subtests in the Matrices Test were analyzed to determine item hierarchical structures for the two groups.

Procedures

A revised method to determine item hierarchies for sets of diehotomously scored items from the framework of ordering theory (Bart and Airasian, 1976) was used to search for item hierarchies in each of the four subtests. Let a success on an item be scored as a 1 and a failure on an item be scored as a 0. An item i is prerequisite to item j if and only if the frequency of 01 item respnse patterns for items i and j respectively is equal to or smaller than some pre-establised error integer value. If the error integer value is greater than 0, then item intransitivities (i.e., if item i is prerequisite to item j and item j is prerequisite to item k, then item i is not prerequisite to item k) are possible. However, transitivity is a basic property of a hierarchy; therefore, a method which allowed no item intransitivities was required. Such a method was developed and used in this study to determine item hierarchies: 1) item response patterns are deleted only if their frequencies are less than or equal to some

tolerance level; 2) inter - item prerequisite relations are derived from the remaining item response patterns with an error integer value of 0. This revised method allows no item intransitivities and includes no infrequently occurring response patterns. Tolerance values of 1, 2, 3, and 4 were used in the deletion of response patterns.

RESULTS

An agreement score was calculated for each twin pair by counting the number of items two twin members agree in their item scores. High mean agreement score was an indication of high heritability of intelligence and a low score would indicate the opposite. The possible range of such scores for 12 items is 1 - 12. The black mean agreement scores were 8.99, 7.99, 7.99, and 8.35 for the four subtests and the white mean agreement scores were 10.09, 8.78, 8.30, and 8.60 for the same subtests. Since these agreement scores were higher for the white twins across subtests, it is inferred that intelligence is slightly more inheritable in the case of whites than blacks, although heritability appears high for both groups.

In another analysis, one member of each black twin pair was randomly selected for one set and one member of each white twin pair was randomly selected for another set. The item hierarchies for these sets of subjects for the four subtests were compared with item hierarchies for the complementary sets of subjects of the same race. No prominent differences were identified between the hierarchies. (In this analysis, four pairs of item hierarchies were assessed for the black group and four pairs were assessed for the white group. In general then, the hierarchical structures generated were quite consistent for both groups.

Another analysis involved the pooling of the black student item response patterns and then those of the white student item response patterns. Item

hierarchies for each of the 12 - item subtests A, B, C, and D were generated for the black and white groups at whole response pattern tolerance levels of 1, 2, 3, and 4. The item hierarchies for tolerance level of 2 are provided in figure 1 - 8.

Enter Figures 1 - 8

As a measure of the structuredness of an item hierarchy, the number of inter-item prerequisite relations was calculated and termed the S - index. Please note that in the extreme case in which there are no inter-item prerequisite relations and thus no item hierarchy to speak of, the structuredness of the item hierarchy is 0. Also note that in the opposite extreme case in which all items are logically equivalent to each other and thus a very dense item hierarchy is indicated, the structuredness of the item hierarchy is n (n-1) where there are n items in the test.

Table 1 depicts the information regarding structuredness and shared structuredness for item hierarchies. The structuredness or S-value for, for example, the black group for the subtest A item hierarchy for tolerance level of 2 is 43. The amount of shared structuredness between the item hierarchies for the two groups for the same tolerance level is labelled as $S_B \wedge S_W$. The amount of structuredness that the item hierarchies for the two groups for the same tolerance level together determine is $S_B \vee S_W$. The amount of structuredness that is shared in relation to the total amount of structuredness determined is labelled $\beta (= S_B \wedge S_W / S_B \vee S_W)$, which provides an index of how similar two item hierarchies are. For tolerance levels of 1 and 2, the item hierarchies for the two racial groups tend to be dissimilar, but, for tolerance levels

of 3 and 4, the item hierarchies tend to be very similar from an examination of Table 1. From an examination of the item hierarchies, it appears that in general the black and white groups engender similar item hierarchies, thus, identifying similar patterns of intellectual skill acquisition and similar instructional sequences for the improvement of intellectual skills, especially for modest tolerance levels of 3 and 4 being used.

However, some items acted quite differently for the two groups. For example, item 12 in subtest A had 11 items as its prerequisites in the white group but only 2 items as its prerequisites in the black group for the tolerance level of 2. Thus for the white group there appears to be a much more extensive chain of intelligence test item skills which must be mastered before item 12 can be mastered than is the case with the black group. Another way to describe this situation is to say that for the black group there may be several paths of acquisition that lead to the mastery of item 12 but for the white group there may be very few paths of acquisition that lead to the same item 12 mastery which would account for the extensive chain of prerequisite items for item 12 in the white group item hierarchy.

In examining the item hierarchies, their structuredness values, and β values, one might consider that much of the variance of those values can be accounted for by the item difficulty ranks and their similarities. Thus one might posit that two item hierarchies for two groups will have a high β value if the ranks of the item difficulties for the two groups are highly correlated. To examine that hypothesis, the Spearman rank correlations relating item difficulties for the black and white groups were calculated and had values of .90, .99, .98, and .98 for the four subtests. The Spearman rank correlations relating those correlations just cited and β values across

the four subtests were calculated and had values of -.25, -.25, -.55, and -.25 for tolerance levels of 1, 2, 3, and 4. Those four negative correlations lend redence to the assertion that the similarity of item hierarchies for two groups cannot be accounted for by the correlation of the item difficulties for the two groups.

DISCUSSION

In this study an ordering theoretic analysis of four subtests of the Raven's Progessive Matrices tests was performed to describe and compare the item hierarchies for samples of black and white students and thus to assess the similarity of patterns of intellectual development and learning hierarchies for intelligence test item skills across the two races. These ordering-theoretic procedures are deterministic and treat error in a non-probabilistic manner; they primarily search for item hierarchies and provide descriptive information regarding item hierarchies and their similarities.

With such analysis hypotheses regarding patterns of intellectual ability acquisition and optimal instructional sequences could be readily formulated and then tested. Also, research to examine whether item hierarchies for different racial, social class, and ethnic groups are due to environmental or genetic factors could be instituted. no assertions are made in this paper as to whether the similarities and differences in the item hierarchies for the two racial groups are due to genetic or environmental factors. In addition, those intelligence test items that tend to be processed differently in different groups can also be identified. Thus ordering-theoretic procedures to analyze item hierarchies are measurement methods that provide information of significance to developmental and differential

psychology as well as to instruction without the socially-divisive statements that have beset the IQ controversy.

References

Airasian, P. and Bart, W. Validating a priori instructional hierarchies. Journal of Educational Measurement, 1975, 12, 163-173.

Bart, W. and Airasian, P. The determination of seven Piagetian tasks by an ordering-theoretic method. Journal of Educational Psychology, 1974, 66, 277-284.

Bart, W. and Airasian, P. An empirical examination of the relationship between test factor structure and test hierarchical structure. Paper presented at the 1976 A.E.R.A. meeting.

Block, N. and Dworkin, G. (Eds.) The IQ controversy. New York: Pantheon, 1976.

Partin, R. The instructional effectiveness of random, logical, and ordering theory generated learning hierarchies. Paper presented at the 1976 A.E.R.A. meeting.

Table 1

Structuredness and Shared Structuredness for Black and White Groups for Each of Four Subtests for Four Different Tolerance Levels

Subset	Tolerance Level	S_B	S_w	$S_B \cap S_w$	$S_B \cup S_w$	$\beta = \frac{S_B \cap S_w}{S_B \cup S_w}$	Percent of Population and Sample Size remaining after Removal of Response Patterns for Different Tolerance Levels: $\%_B$ (N_B)	$\%_w$ (N_w)
A	1	24	30	22	32	0.68	81%(293)	90%(400)
	2	43	54	33	64	0.52	74%(267)	84%(374)
	3	53	60	52	61	0.85	65%(231)	80%(359)
	4	65	76	60	81	0.74	60%(217)	75%(335)
B	1	6	14	6	14	0.43	63%(228)	75%(334)
	2	41	51	23	69	0.33	46%(166)	65%(290)
	3	55	53	44	64	0.69	38%(136)	57%(254)
	4	68	56	51	73	0.70	31%(112)	53%(238)
C	1	1	5	0	6	0.00	61%(221)	71%(317)
	2	10	34	8	36	0.22	47%(169)	55%(247)
	3	40	44	30	54	0.55	32%(115)	46%(205)
	4	62	57	47	72	0.65	24%(87)	41%(185)
D	1	3	18	2	19	0.11	70%(255)	78%(350)
	2	27	38	15	50	0.30	58%(209)	65%(290)
	3	57	48	48	57	0.84	47%(170)	61%(272)
	4	62	61	50	73	0.69	43%(154)	57%(256)

Figure Captions

Figure 1. Ordering theoretic item hierarchy for white twins (n = 374, % N = 84%, S = 54) for Raven's Matrices subtest A for tolerance level of 2.

Figure 2. Ordering theoratic item hierarchy for white twins (n = 290, % N = 65%, S = 51) for Raven's Matrices subtest B for tolerance level of 2.

Figure 3. Ordering theoretic item hierarchy for white twins (n = 247, % N = 55%, S = 34) for Raven's Matrices subtest C for tolerance level of 2.

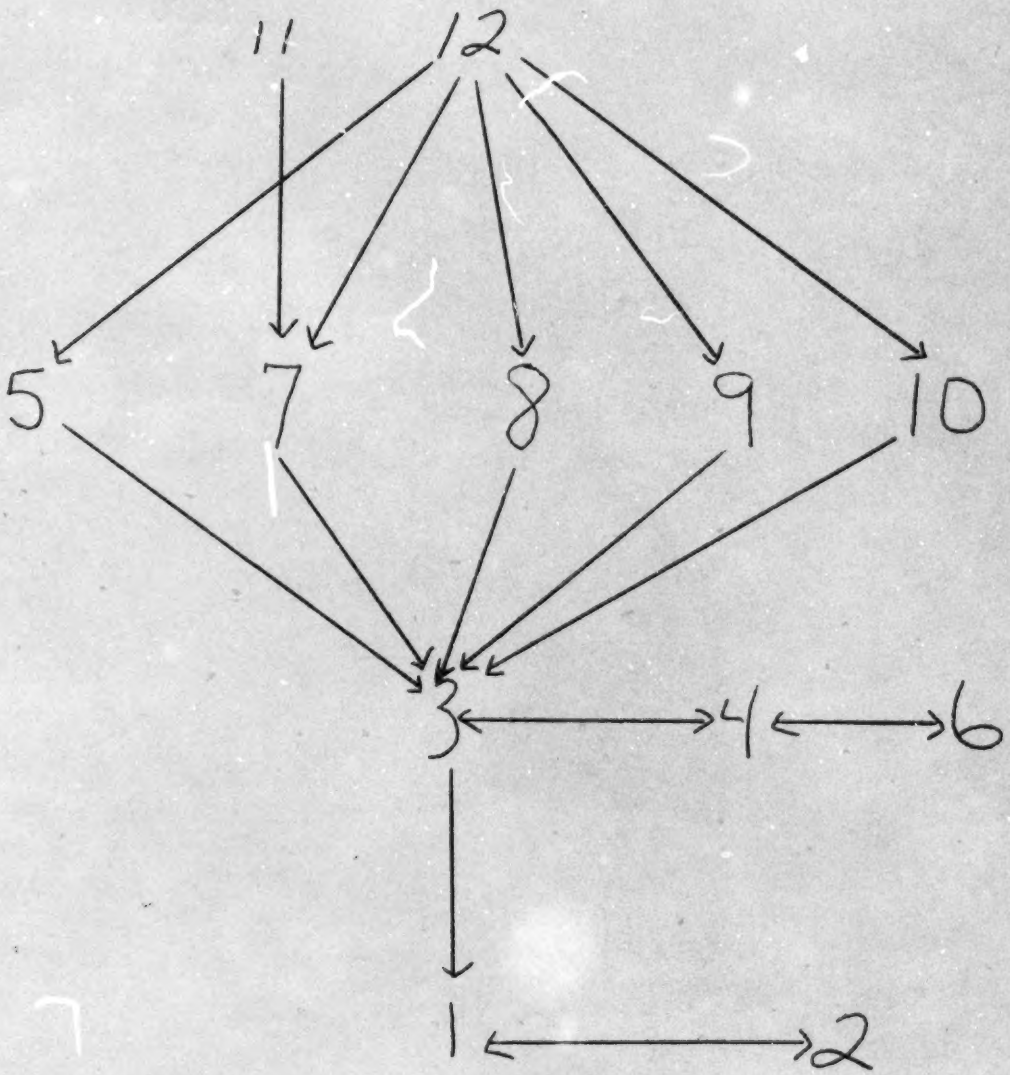
Figure 4. Ordering theoretic item hierarchy for white twins (n = 290, % N = 65%, S = 38) for Raven's Matrices subtest D for tolerance level of 2.

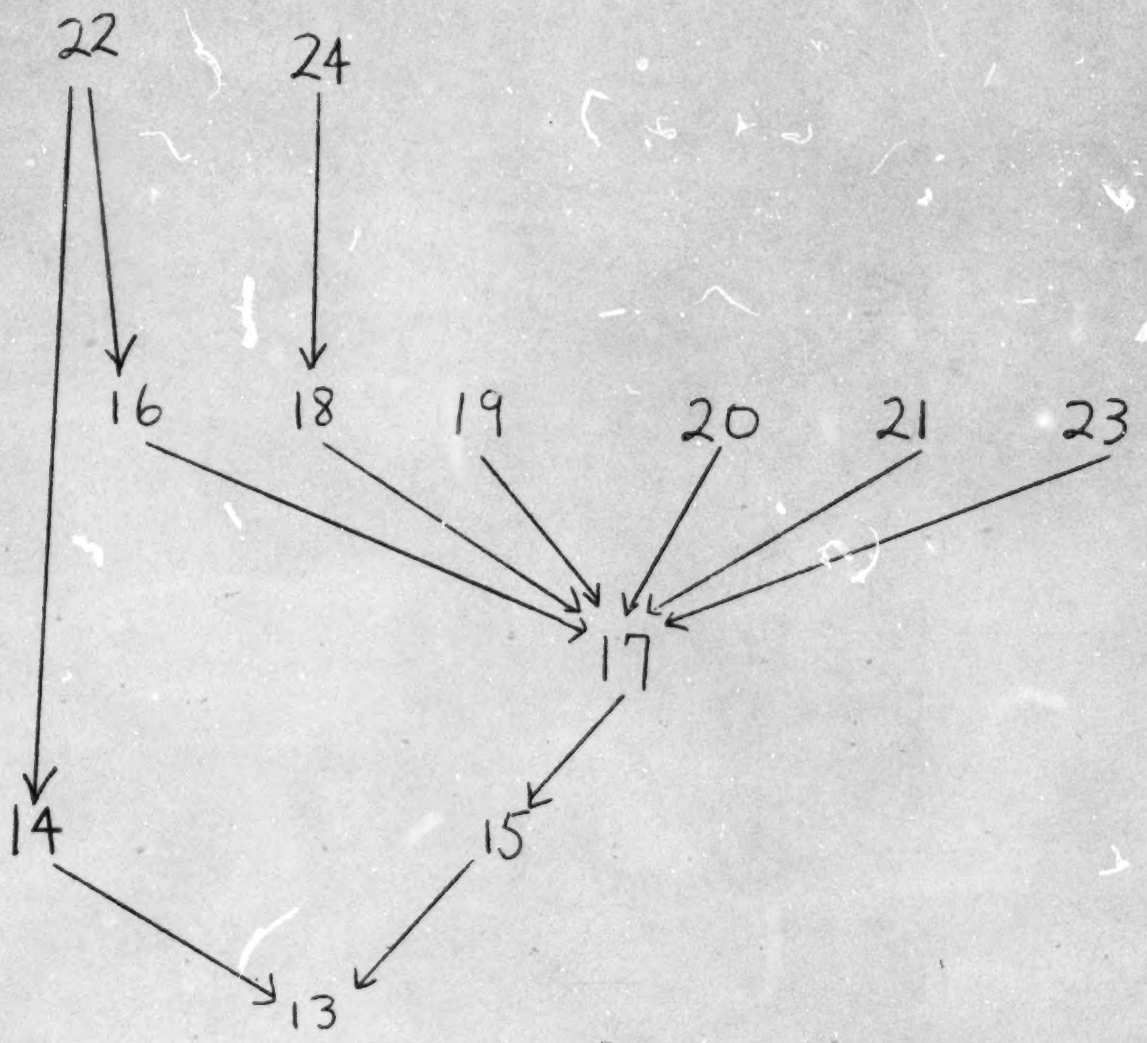
Fibure 5. Ordering theoretic item hierarchy for black twins (n = 267, % N = 74%, S = 43) for Raven's Matrices subtest A for tolerance level of 2.

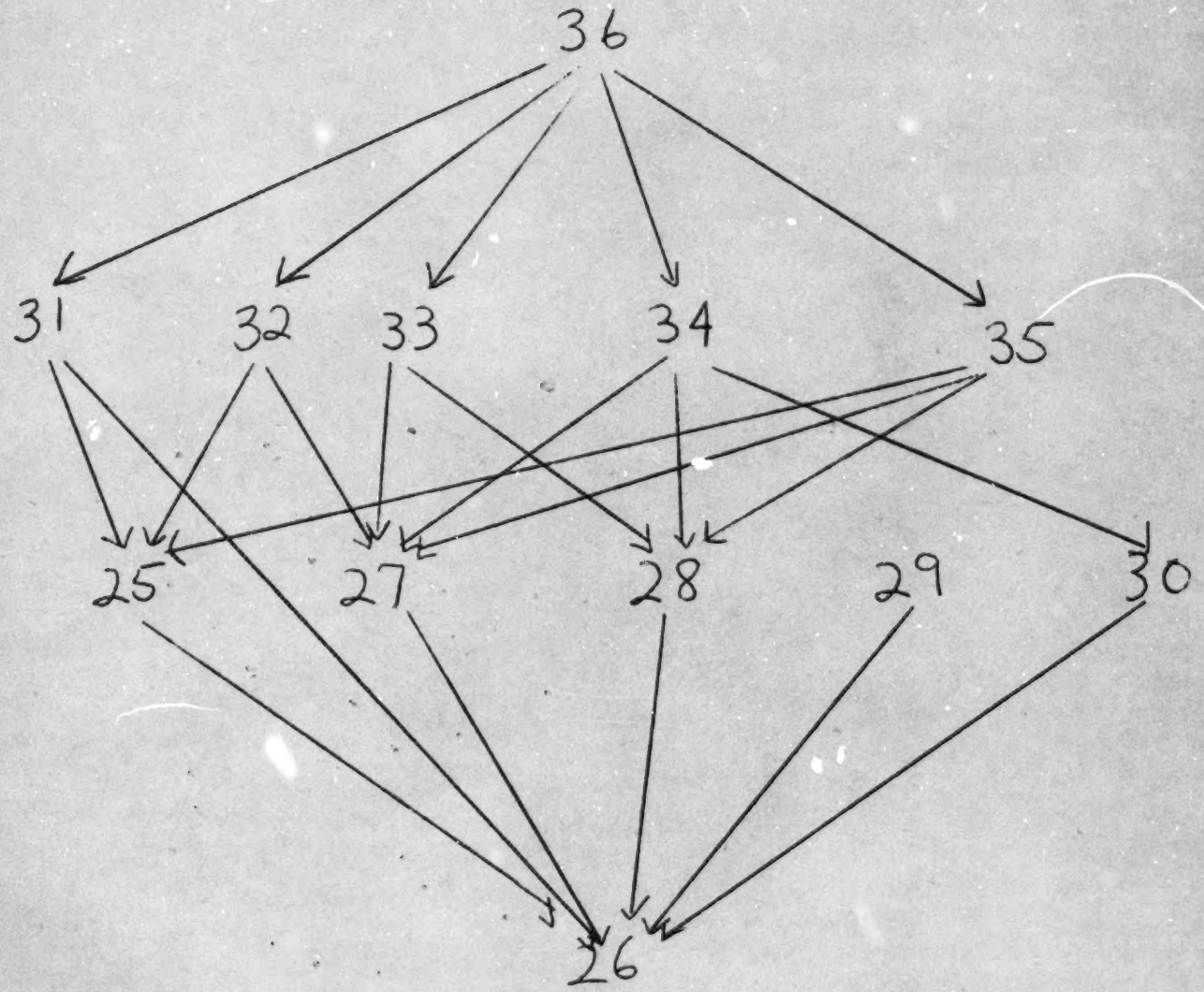
Figure 6. Ordering theoretic item hierarchy for black twins (n = 166, % N = 46%, S = 41) for Raven's Matrices subtest B for tolerance level of 2.

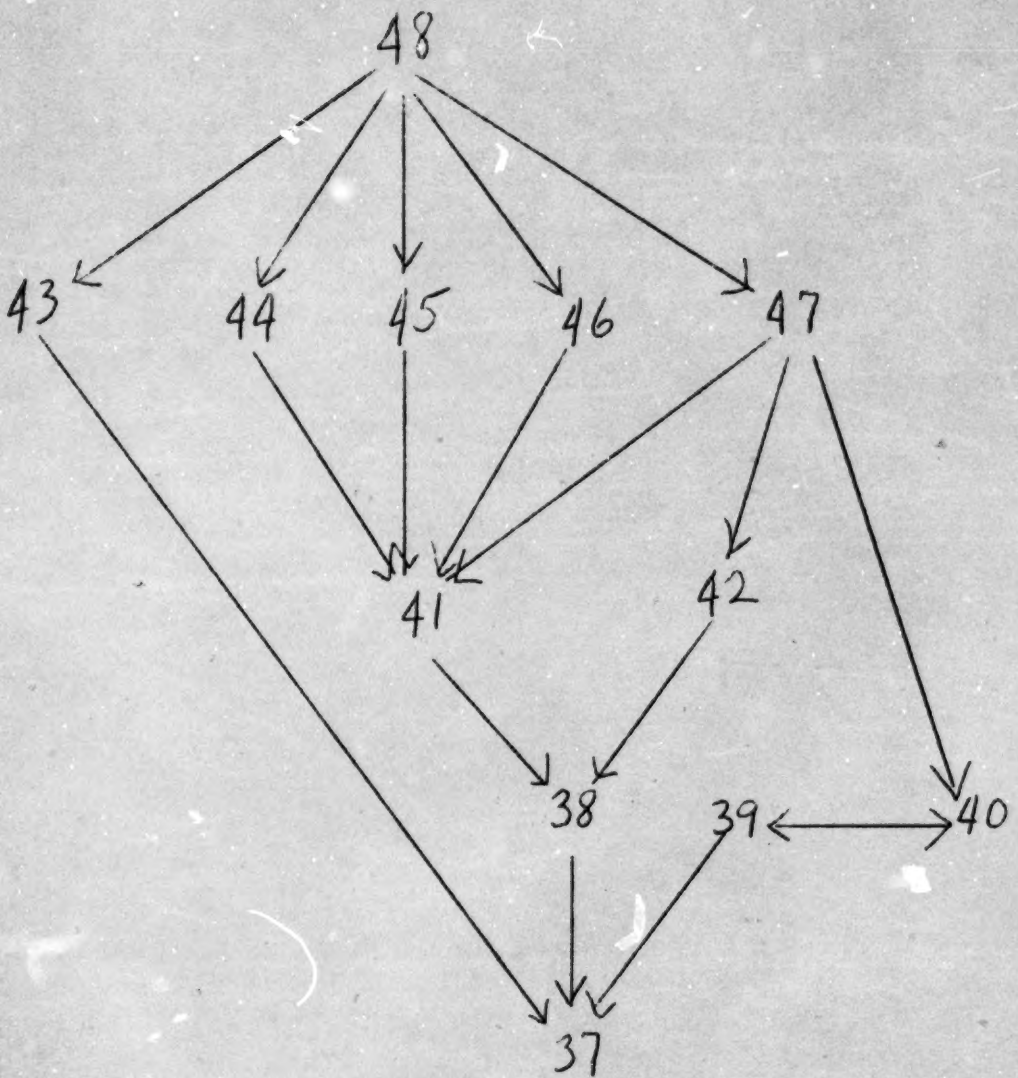
Figure 7. Ordering theoretic item hierarchy for black twins (n = 169, % N = 47%, S = 10) for Raven's Matrices subtest C for tolerance level of 2.

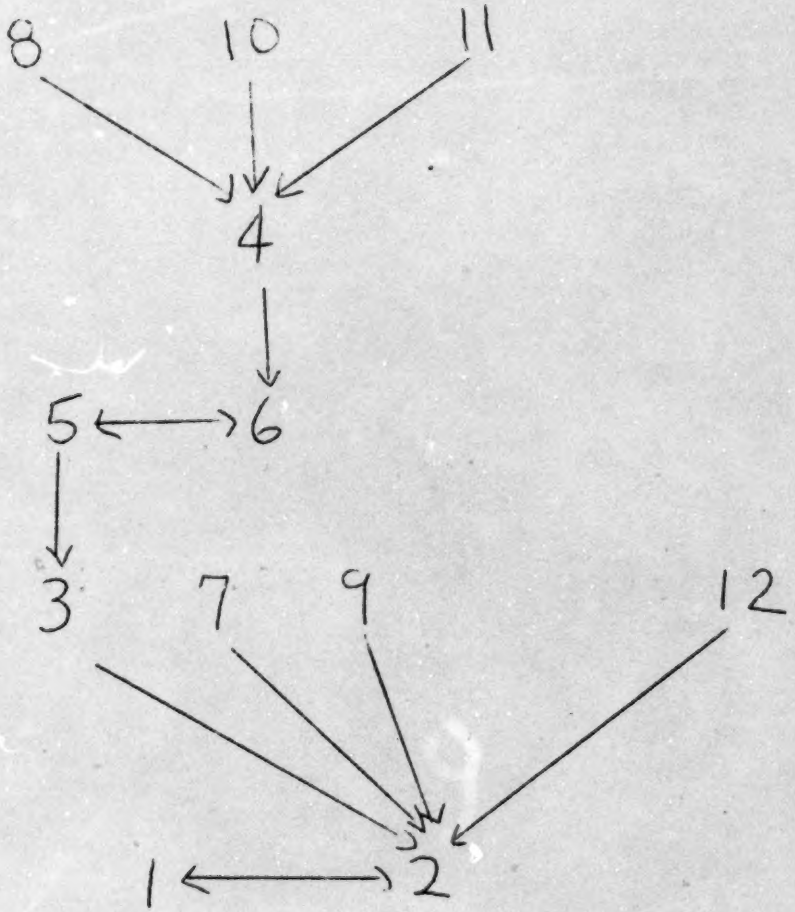
Figure 8. Oredring theoretic item hierarchy for black twins (n = 209, % N = 71%, S = 27) for Raven's Matrices subtest D for tolerance level of 2.

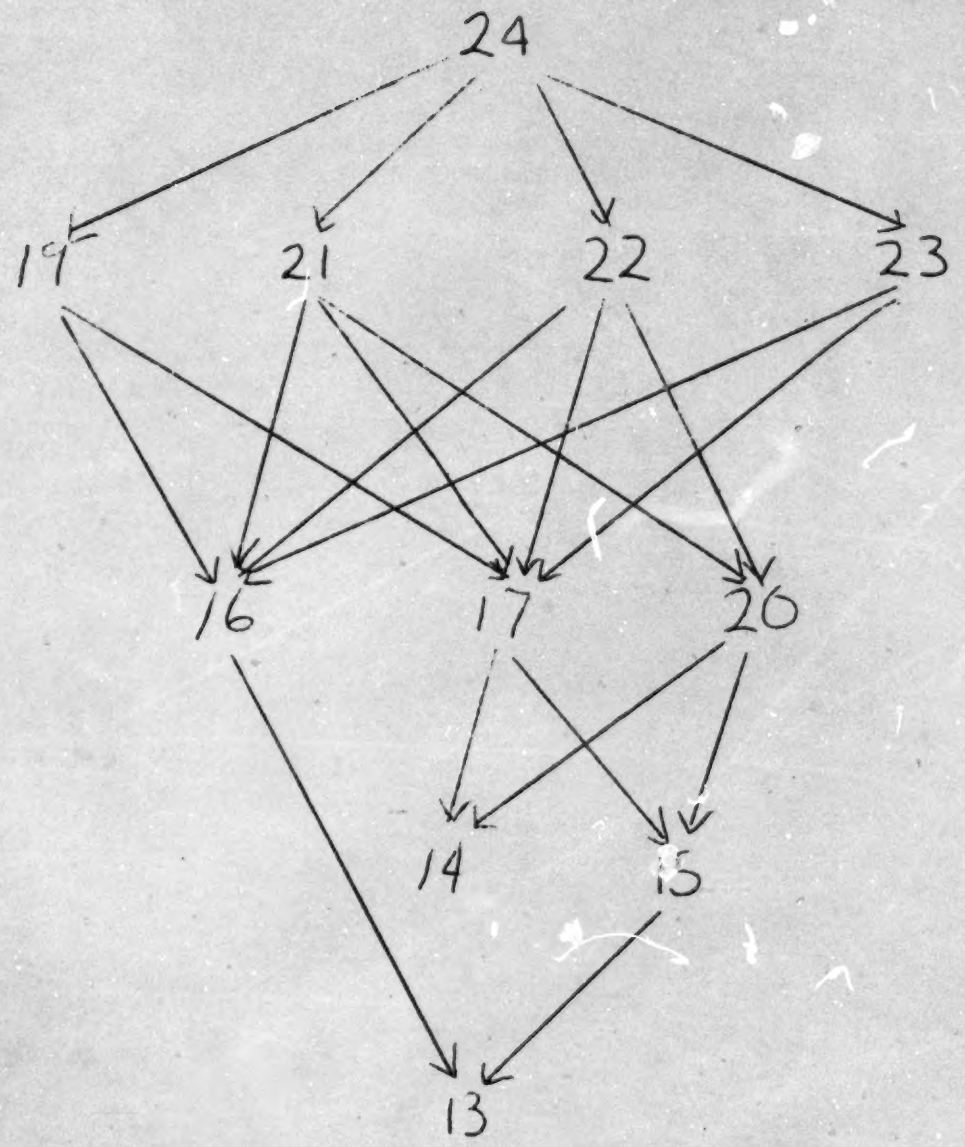

11
12
5
7
8
9
10
3
4
6
1
2

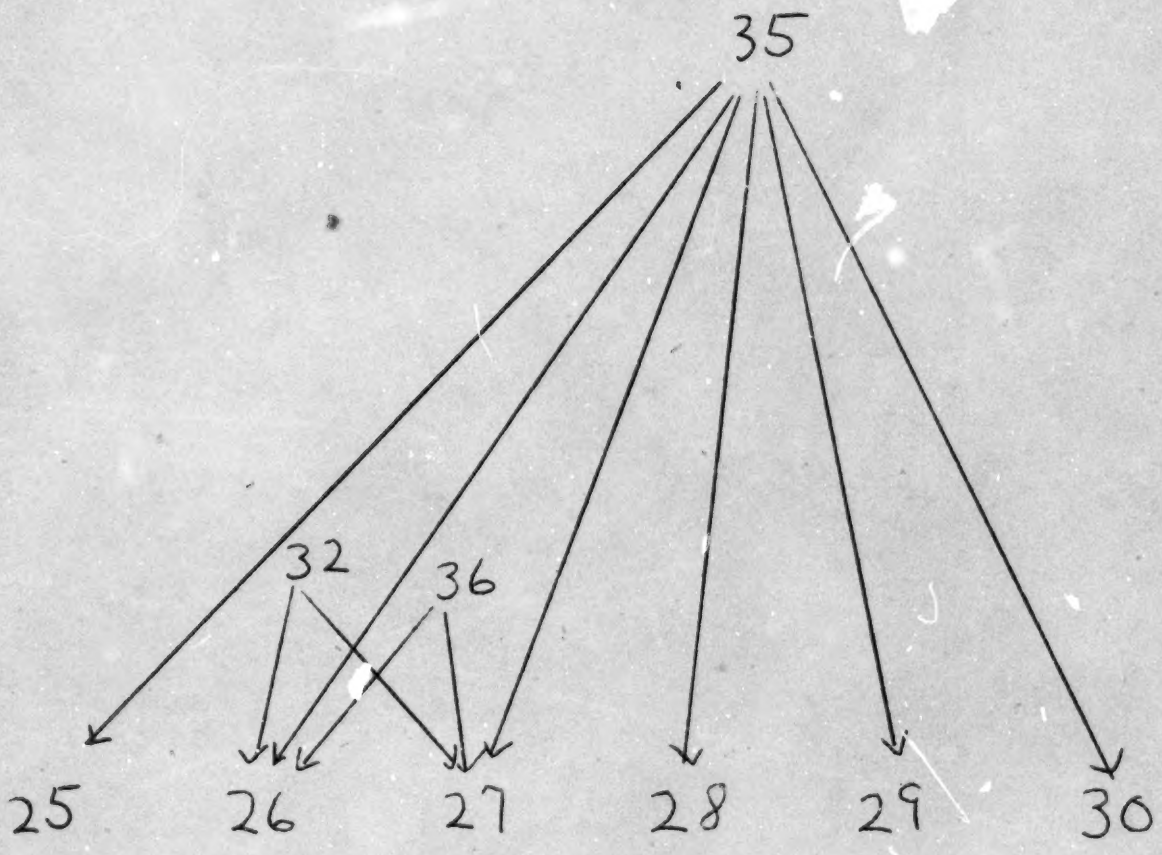

22
24
16
18
19
20
21
23
17
14
15
13

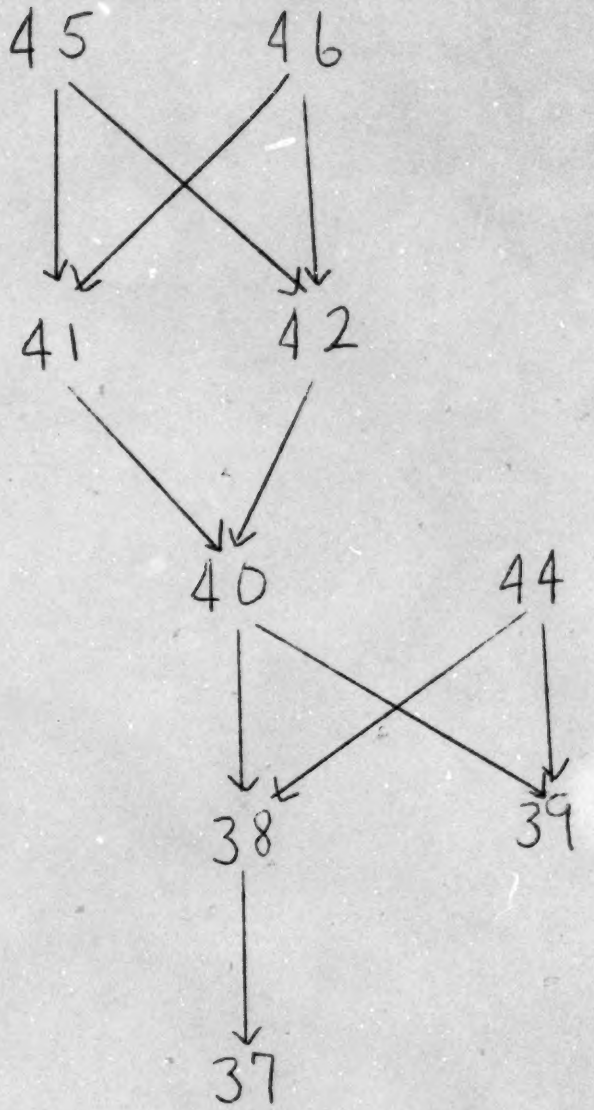

36
31
32
33
34
35
25
27
28
29
30
26


48
43
44
45
46
47
41
42
38
39
40
37


8
10
11
4
5
6
3
7
9
12
1
2


24
19
21
22
23
16
17
20
14
15
13


35
32
36
25
26
27
28
29
30


45
46
41
42
40
44
38
39
37